Australian Government
Department of Defence
Defence Science and Technology Group

Trusted Autonomy:

Conceptual Developments in Technology Foresight

Scott Wheeler

Joint and Operations Analysis Division
Defence Science and Technology Group

DST-Group-TR-3153

ABSTRACT

This report provides the synopsis of a five year collaborative program of research between DST Group and the University of South Australia in the study of autonomous systems concepts. The purpose of the program is to establish a methodological means of technology foresight, to assess how future technologies shape or contribute to performance in autonomous systems and to identify key technologies of greatest impact. We propose a new model for the categorisation and assessment of autonomy, which provides a systematic and auditable way to explore emerging technologies for autonomous system by means of parametric investigation. Outcomes might then inform decision makers in Defence and National Security to shape the future of policy, strategy, emerging concepts, and force development in autonomous systems and their related technologies.

Published by

Joint and Operations Analysis Division
Defence Science and Technology Group
506 Lorimer St
Fishermans Bend Victoria 3207 Australia

Telephone: 1300 333 362

AR-016-395
September 2015

APPROVED FOR PUBLIC RELEASE

Trusted Autonomy:

Conceptual Developments in Technology Foresight

Executive Summary

This report provides the synopsis of a five year collaborative program of research between DST Group and the University of South Australia in the study of autonomous systems concepts. The purpose of the program is to establish a methodological means of technology foresight, to assess how future technologies shape or contribute to performance in autonomous systems and to identify key technologies of greatest impact. This program describes the connection between:

a. current and future technologies for trusted autonomous systems;

b. predictions about how those technologies evolve or emerge over time; and

c. how those technologies are brought together to generate operation capability.

This is, in simple terms, what we mean by technology foresight. Technology Foresight is the field of scientific regard which investigates the emergence, performance, or impact of technology across society. It aims to describe usage or uptake, and evolving trends, in technological development over time. Through doing so, it seeks to determine the implications of those developments both current and forthcoming.

This report provides a means by which the foresight process can be conducted, for the study of autonomy in engineered systems, and the record of the work that has led to its development. The approach is based on a hierarchy of four chosen factors:

1. technical specification of the system

2. mission complexity

3. context of employment

4. trust in the system.

Each of these factors is decomposed into their most significant constituent elements to define a new conceptual model for autonomy. This model is to be employed as part of a technology foresight process and its outcomes will inform a range of stakeholders in the autonomy, autonomous systems and automation client communities.

Author

Dr Scott Wheeler
Joint and Operations Analysis Division, DST Group

Scott Wheeler joined DST Group as a Research Scientist after completing a PhD in mathematics at the University of Adelaide. Scott previously managed the Complex Adaptive Systems Task in Land Operations Division at the Edinburgh site before moving to the Defence Systems Analysis Division in Canberra's Russell Offices. In Russell, Scott worked in the domain of Capability Analysis and was the Science Advisor to the Missile Defence Coordination Office. More recently he represented Australia as the National Lead for TTCP AG-14 Complex Adaptive Systems. Scott was previously a Visiting Research Fellow at the University of NSW, Australian Defence Force Academy and now works for the Joint and Operations Analysis Division in Fairbairn, Canberra.

Acknowledgements

Thanks are extended to all of those who contributed to this report. There have been almost too many to mention individually. Further thanks are specifically extended to staff within the University of South Australia, Dr Anthony Finn and Dr Brian Mekdeci, for their expertise and research (Finn, 2011; Finn, 2012; and Finn & Mekdeci, 2014)) which have been invaluable to the development of this report. References to their work have been included throughout the report.

Contents

List of Tables

List of Figures

Acronyms

Acronym	Definition
ADO	Australian Defence Organisation
DDR&E	Department of Defence Research and Engineering
DST Group	Defence Science and Technology Group
TFF	Technology Forecasting and Futures
US	United States

1. Introduction

1.1 Background

Technology continues to develop; a seemingly certain feature of the modern era and one of its greatest challenges. Defence organisations across the world are increasingly being pressed to employ a range of disparate technologies in new and innovative ways to retain a persistent capability advantage. Those that embrace and exploit those new technologies in a considered manner are better positioned to adapt and respond to emerging threats, develop advanced capabilities, evolve their operating concepts, and shape their force structure.

The Australian Defence Organisation (ADO) looks to the Defence Science and Technology Group (DST Group) to lead the identification and assessment of emerging technologies that offer advantages or represent threats to operations. Technologies for autonomous systems are important to Defence because they have the potential to extend the reach and capability of traditional forces while reducing operational footprint and threat to personnel. This increases access to regions of operational interest, especially in contested regions, possibly enhancing control, freedom of manoeuvre, and denying the area to adversary forces. In the longer term, autonomous systems may also reduce the cost of operations, capability acquisition and training (Ellis et al, 2005).

1.2 Program Objectives

Technology Futures and Forecasting (TFF) group, within the DST Group, is a collaborative research facility for the study of emerging and disruptive technologies. It encourages participation and understanding of related issues across academia, government and industry through strategic alliances. Leveraging from community engagement, it identifies areas of threat and opportunity in developing technologies and provides foresight to policy, strategy and capability development for the ADO and its strategic partners.

As a part of its program, TFF participates in a number of bilateral and multinational programs. Through mutual agreement, these arrangements provided a formal mechanism to execute a number of focused studies on key technologies of interest for benefit to all members. Under this mandate, technologies for autonomous systems have been assessed as critical to Australia because of their impact on current operations and anticipated influence in the future battle-space as a force multiplier (see for example DSTO, 2015). Additionally, the related technologies are becoming widely available and at lower cost. This trend presents a high risk to Defence, whereby the technologies might be adopted by a range of interested parties, to generate disruptive effects. In response, the TFF stood up their Trusted Autonomy program (see Appendix A), established in 2007. Since then, TFF has explored a number of themes in Autonomy in collaboration with centres of excellence in academia including:

- Robotic Systems (Zimmer, 2007);

- Robotics for Hazardous Environments
  (Trevelyan, 2007; Trevelyan et al, 2007; Cebon & Samson, 2008);

- Micro-Systems (Spinks & Alici, 2008); and

- Trusted Autonomy (Finn, 2011).

These case studies each delivered a comprehensive technology review within their focused field of regard. In these early studies, the objective was to review the characteristics and technical properties of the technologies and systems used to develop autonomous capabilities. Having established a broad understanding of the technology domain and its applications, TFF now seeks to understand the implications of technologies for Trusted Autonomy on Defence and National Security operations. Hence, the focus has moved on from investigating and reporting about technologies and related systems for automation towards exploring the potential future capabilities which might be realised and their utility.

In executing this new program, TFF established a Research Agreement with the Defence and Systems Institute (DASI) at the University of South Australia (UniSA) on 16-May 2011.[1] Working in conjunction with UniSA then provides the TFF a unique opportunity to leverage from their academic program[2] and grants access to a wider range of new and developing technologies for autonomy. This also generates synergies within the program in DST Group, including the Autonomous Systems Initiative. In this respect, the research undertaken does not duplicate effort but is carefully directed towards those issues which add-value, by developing or progressing knowledge about Autonomous Systems and addressing the requirements of the Defence community, including the Project Arrangement, principle clients in the ADO and significant stakeholders.[3]

In April 2012, UniSA delivered the compendium report *Implications for Autonomy & Autonomous Systems* (Finn, 2012).[4] This report contained a breakdown of Autonomous Systems into their components and technology areas. Using this approach, UniSA then explored the potential role for those technologies, corresponding functional requirements, and implementation challenges. Our current work takes a complementary approach to

---

[1] This agreement was administered under the Strategic Alliance with the Defence Systems Innovation Centre, established with DST Group through Joint Venture Agreement on 2-March 2009.
[2] Research at UniSA includes: UAV-based atmospheric tomography; acoustic technology for UAVs operating in civilian airspace; modelling of autonomous and robotic systems; collaborative control and multi-robot coordination; and legal and ethical considerations in autonomous decision-making.
[3] This includes the exploitation of products, to operationalise and institutionalise research outputs, across (1) the five-eyes intelligence community under the Emerging Technology Analytical Panel which operates under the Quinquepartite Technical Intelligence Steering Panel, and (2) the five-eyes science and technology community under the Emerging and Disruptive Technology Action Group under The Technical Cooperation Program. Australian clients include the groups of Capability Development, Vice Chief of Defence Force, Strategy, and all three of the services.
[4] A compilation of knowledge sourced from experts in the University of Sydney, University of New South Wales, University of Technology Sydney, Queensland University of Technology, Australian National University, UniSA, and DST Group.

extend this study. While the 2012 report focused on documenting technology for Autonomy and trends in its development, this new work instead presents a pragmatic approach which could be employed for the purposes of technology scanning and foresight (Finn & Mekdeci, 2014). That is, this work focuses on presenting concepts, models and frameworks to support decision-makers and capability developers and is not concerned about particular technologies in themselves.

1.3 Scope

This report focuses on the study of applications of TAS for Defence and National Security, enabling technology and its integration, technology forecasting, and operational implications. In order to make meaningful insights, it is first critically important to understand what these systems are and how they work. This understanding permits us to make judgements about their capabilities into the future and the likely implications of those capabilities to operations. Two research questions are then important:

1. *What does it mean for an engineered system to be autonomous, so that decision makers have a reasoned understanding of their capability and can make informed decisions about their potential applications and operational impact?*

2. *What are the criteria by which autonomy can be assessed, so that decision makers can make rational assessments about the effects of technology changes on the capability and value of autonomous systems?*

In answering these questions, we establish the baseline against which we can assess how technologies shape or contribute to autonomy and to identify those key technologies with greatest impact. The ultimate aim is then to determine the technologies which affect autonomy in engineered systems; such that, the change in capability of those systems is significant in foreseeable future operations. In this sense, this study seeks to describe the connection between:

a. current and future technologies for trusted autonomous systems;

b. predictions about how those technologies evolve or emerge over time; and

c. how those technologies are brought together to generate operation capability.

Related technologies could be said to be on a critical development path for Autonomous Systems. Identifying possible evolving trends in relevant technologies, the capability of those systems, and their operational implications is then one of the greatest objectives of technology scanning and technical foresight. However, the future is rarely predictable. For practical purposes, the foresight process seeks only that decisions made under uncertainty are more informed, better argued or considered. Outcomes of this work, and related initiatives, then support the Australian and allied capacity to improve reliability and safety in complex and dynamic environments which are shared by personnel, property and facilities, and other systems. The work ensures that Defence' future capability is safe and

fit for purpose, and that Australia is a smart buyer of autonomous systems and their technologies.

2. Trusted Autonomy

2.1 An Introduction to Trusted Autonomy

Attempts at describing Trusted Autonomy within the literature are many and varied. To appreciate this, we only have to consider that the term itself has only recently been coined. Over the last 50 years, the technology of 'autonomy' has evolved from human-operated mechanical systems (under direct manual control), to human-supervised automated systems (requiring human-in-the-loop interaction), to unsupervised automatic systems (and direct execution without human oversight) (Lomuscio, 2015). Thus, definitions have changed significantly over time and recent attempts have struggled to entirely capture all facets of autonomy. Efforts towards defining the term have been further confounded by domain of application, be it software systems, engineering, jurisprudence, or even philosophy all have alternative competing perspectives.

Put in simple terms, *trustworthiness* is how well a system performs a stated task without operator intervention. *Trust* then relates to the operator's perception of a systems capability in fulfilling functions or tasks. Trust in autonomy is then a multi-dimensional construct influenced by expression of purpose, intention and role; approaches to developing and determining trust; functional aspects such as system capability; and system reliability within an operating environment. This complexity is difficult to capture in a single concise definition of the term because any definition must be sufficiently rich as to capture all of these ideas. We devote Section 3 of this report to exploring how the literature has approached this problem.

However, we also wish to make some practical headway, towards developing a simple and short definition of the term. Towards this purpose, we adopt the position that:

> *"all systems which may be described as 'automated' or 'autonomous' are still merely machines, constructed, or programmed, to relieve a human operator of some decisions and actions that would otherwise have been done manually"*
>
> (McFarland, 2015).

This is inherently useful, because in law there is little distinction between the technical operations of autonomous and automated systems. Ultimately, the same legal principles applicable to the employment of all systems in Defence and National Security operations also apply to the employment of Trusted Autonomy and remain the only clear, meaningful, and lawful basis for use. Technical specifications of a trusted system become the measure by which operators delegate control - namely, the delegation of decisions from the user to the system. Delegation of control is important because of the potential transfer of oversight, responsibility and, ultimately, liability in the event of a serious violation of the law. But any transfer of responsibility does not reside with the system itself. Instead, lawyers are grappling with concepts of system designer/developer responsibility. Defining Trusted Autonomy in a system includes a process of defining lawful regulation of the system.

Simplifying this concept leads us to propose the following definition for Trusted Autonomy, namely being:

> *Trusted Autonomy is an emerging field of research, with the aim to minimise operator intervention and oversight in the planning and execution of tasks conducted by autonomous systems.*
>
> (Wheeler, 2015).

The key point here is that the definition focuses on the objective of the field of research and not the function of the system. For the remainder of this report, we assume this as our working definition and leave the detailed exploration of the many faceted concepts in Trusted Autonomy until Section 3.

2.2 Benefits of Autonomous Systems

Militaries, governments and industries around the world are becoming increasingly interested in autonomous systems, as their technical capabilities improve and their entry-level unit cost decreases. However, the demands placed on autonomous systems are high; they need to perform complex tasks accurately, efficiently and within unusual or dangerous environments. To support research, development, acquisition and investment, there is a need to understand how advances in technical capability will underpin the advancement of capability in autonomous systems and to understand the impact those advances may have in an operational context. Compliance with legal and policy frameworks, systems certification, and cultural acceptance are also constraint to the adoption of autonomous technologies. Policy makers need to understand what capabilities autonomous systems will have in the future, so that the can legislate and plan accordingly. It is then no surprise, the development of autonomous systems and integration of those systems into operations is challenging. Tasks are often poorly understood and ill-defined, with ambiguous goals that are subjective or contextual. Even when tasks are well defined and the goals are clear, they can be extremely difficult for an autonomous system to execute.

Notwithstanding, increased use of autonomous systems presents the possibility of enormous gains. Proponents of autonomous technology are excited by the potential to perform tasks beyond what is currently achievable by humans alone. Machines can perform some operations far quicker, and with far fewer (if any) errors than humans can, making certain tasks in areas such as accounting and data analysis, trivial. Machines are not susceptible to factors that affect human performance, such as fatigue and emotion, thereby making them more reliable than humans could ever be in some tasks. Automation offers many advantages through their computation processes alone:

- Machines are not (directly) subject to limitations in human thought process, bias, experience, or capability.

- Direct computation can be used to explore and evaluate large numbers of possible states of interest, far more that a human could in the same period.

- Autonomous systems models operate, without loss of objectivity, in contexts where complexity would degrade human performance.

- The architecture may permit the machine to execute simulations to identify patterns or potentialities, which might otherwise be unpredicted or unanticipated.

- The logic, or thinking, or an automated system can be traced, and although this may be stochastic it can still be audited.

The second core benefit of autonomy resides in the physical characteristics of the system themselves. Autonomy offers the possibility to liberate personnel from dull, dirty or dangerous work, and to complete physically demanding tasks that are currently out of human reach (Sharkey, 2008). Autonomous Systems may be able to reduce physical workloads, intervene in hazardous or life-critical environments, improve personal well-being, security, and benefit commercial and government enterprises through: [5]

- **Increased Productivity**. Correctly introduced, automation increases productivity while maintaining quality. This facilitates a faster operational tempo or production cycle, with greater efficiency and improved reliability.

- **Cost Efficiencies.** Automation also permits 'lean' manufacturing processes. Automation simplifies labour-intensive tasks, which can lead to a reduction in workforce costs, and can reduce generation of waste materials and products.

- **Improved Quality**. Automated systems provide consistent results, within specification, thereby eliminating quality control errors associated with human error. Processes and tasks can then be carefully controlled and regulated so that the outcome is consistent or more reliable.

- **Better Safety**. Robots are able to endure hazardous environments, removing personnel from settings which would otherwise place them at risk. Some tasks that machines can perform are impossible for humans, particularly those with environmental constraints to life-support. Machines could perform tasks in space, deep-sea or deep-earth that are well beyond the range of human interaction. Other possibilities include frigid, explosive, toxic, foundry, cleanroom or other environments hazardous to human health. Similarly, systems used in food, medical, or pharmaceutical applications reduce opportunities for contamination.

- **Resilient to Failure**. If a system can perform a complex task without human intervention, then it can do so remotely and without the need for communication. In addition to being less vulnerable to human error, the system will also be less susceptible to communication failure, denial or interception.

In other words, automation is ideally suited to do the work that humans do not (or should not) want to do. This frees the workforce to focus their efforts on other tasks while also

[5] These benefits are cited from Finn (2012, p.31) with minor changes.

reducing risk to personnel and distancing them from harm. This makes economic sense, as human resources are expensive and subject to high liabilities. Operational 'footprint', being the presence of personnel in theatre, can also be reduced. Automation can act as a force multiplier. If tasks that currently require humans to perform can be executed safely and effectively by machines, then those humans can be used to do other tasks, thereby increasing the overall size and effectiveness of personnel.[6]

Automation is not without its disadvantages. While automated systems perform well in controlled environments and repetitive tasks, they are notoriously inflexible when dealing with uncertain environments, and new or novel situations. Humans may make mistakes on routine tasks, but they often excel at recognizing and handling exceptional situations. Automation arguably also reduces interpersonal communication and socialisation, within communities who are extensive users or beneficiaries of autonomous services. There is also a cost, in degradation of fundamental competencies and increased dependency on technology. Both proponents and opponents of autonomous technology can agree, the implications (possibilities and repercussions) of emerging autonomous technologies is not fully understood.

2.3 Drivers and Applications

Autonomous capabilities have grown with the times, benefitting from miniaturisation, increases in computational power, and efficient power technologies. What is more significant is the perception of trust within society is itself also rapidly changing. Autonomous systems are now able to perform a range of tasks more proficiently than their human counterparts, more economically and with greater reliability in service. Lower order intelligence is also embedded, imperceptibly and pervasively into almost all modern Information Technology applications on infrastructure we take for granted every day.

Systems will undergo continual development and evolutionary advancements over time, given the considerable commercial interest. However, the commercial environment is comparatively benign and Australia will need to invest in this capability to secure its systems against active threats. This investment will need to occur across all levels, from platform and sub-system through to national infrastructure. Due to the scale and cost of this endeavor, industry partnerships are likely to become increasingly important. The regulatory environment for the use of autonomous systems is also likely to grow more restrictive as the capabilities of autonomous systems increase. This may present a challenge because trust is an essential element for its application and Australia will need to

---

[6] The vast majority of autonomous systems currently in existence, or in the near future, are supported by or interface with human operators. Many high-profile systems do not demonstrably result in a net saving in workforce. For example, the Chief of the US Air Force, General Schwartz, was quoted as saying that the Predator MQ-1, one of the world's most famous and successful unmanned aerial vehicles was "manpower-intensive", requiring 189 support personnel for a single orbit in Iraq or Afghanistan (Pocock, 2010). However, these instances are not supportive of general trends. Crew numbers in defence, particularly in maritime vessels which are limited in space, are decreasing as autonomy is integrated into control systems, and with clear benefit (Scofield, 2006). Further evidence may be sought from industry, engineering, and commerce which have only shown reduction in workforce numbers where systems or process automation has been introduced.

have overcome legal, cultural, and ethical obstacles to the application of autonomy. Adversaries from different backgrounds may not be subject to the same constraints or imposed boundaries.

Aside from operations in the most structured and benign of environments, we are still a long way from being able to automate many of the relevant processes to a level where no human attention or intervention is required. Current technologies excel at making decisions within a confined context but 'intelligence' in computing has not fundamentally advanced despite impressive advances in technologies for computation. Enough progress has been made, however, for the commercial development of a wide variety of applications and products that require little human oversight. On the basis of commercial returns from early applications, cost reductions in key technologies are being derived from ever more sophisticated levels of autonomy.

Nevertheless, significant investment is still required. While the challenge in the long-term is primarily technological, the challenge in the short-term is one of investing in the science of developing requirements and determining value propositions, developing efficient and effective systems engineering processes to deliver robust manufacturing and schedule predictions, determining how best to integrate such solutions into current or adapted societal processes and organisational constructs, and transitioning technology into a product.

In this regard, despite a recent explosion in the commercial arena, it seems to be generally accepted that widespread application of solutions incorporating full-scale, general autonomy is a decade or more away. General autonomy can only be realised through a breakthrough in the science of artificial intelligence and this is not perceived as likely in the short term. Nonetheless, as many of the lower level technologies have advanced to the point where they may now be employed in an increasing number of semi-autonomous applications that are economically viable, practical, and that provide real value, a range of commercial products are emerging. Analysis of civil markets indicates the commercial-off-the-shelf products are driving consumption, across casual and low-cost markets and high-end investment alike. In this, the following sectors most likely to benefit from commercial development (EUROP, 2009; MASSTLC, 2009):[7]

- **Automotive & Transportation**. Robotics technology is already appearing in the form of advanced driver assistance and collision avoidance systems. Public transportation is another area that is expected to become increasingly automated. As robotics technology continues to improve and mature, unmanned transportation systems and solutions developed for limited scale environments such as airports will be adapted for implementation in urban centers and other general purpose environments.

- **Energy & Environment**. The emergence of robotics technology applications, especially in the areas of automating the acquisition of energy and monitoring the environment, presents significant commercial opportunity and environmental

[7] Cited from Finn (2012, pp. 50-53), as were the list of beneficiary sectors, with minor changes.

benefit. Autonomous and unmanned systems are currently used in operations that are inaccessible or hazardous to people. Examples include exploration of collapsed buildings, sewer inspection, and examination of underwater pipelines. At present these robots are almost always tele-operated. Future systems will require much higher levels of autonomy with faultless operation in environments where communications may be restricted and task imperatives are time-critical.

- **Entertainment & Education**. This area, perhaps more than any other has seen the early emergence of robotics technology enabled products. Robotics provides students with a compelling and tactile avenue to learn key mathematical and scientific fundamentals. Motion simulators, robotics teachers, guides, sports trainers, toys, companions for the elderly are possibilities. Keys to success will be multi-modal communication; not just of the robot's state to their human companions, but assessment by the robot of the human's emotional and physical state or intent. Delivery to the mass market at a level of functionality and at a competitive price sufficient to generate interest is also a major challenge.

- **Healthcare & Quality of Life**. The current application of robotics technology to provide tele-operated surgical systems represents only the tip of the iceberg. The technology holds potential to help control costs, empower healthcare workers, and enable aging citizens to live longer in their homes. Robotic assistants will require an intimate level of interaction and compatibility if safe and dependable operations are to be engendered be it in the workplace, in public, or at home. Their operation will almost certainly aspire to individual tasks or entire sequences of tasks being undertaken autonomously. However, they will almost certainly require a degree of manual intervention.

- **Manufacturing & Logistics**. Beyond the traditional application of robotics technology to automate certain assembly line functions, there is tremendous potential to further automate the manufacture and movement of goods. In particular, the technology promises to transform small scale manufacturing operations and aid the transition of manufacturing back to western economies. Existing industrial robotic systems are typically set to work on a single operation or task for long periods of time. High labor costs and a shortage of skilled laborers will increasingly put pressure on industry to adopt robotic work practices that make use of advanced systems that are able to cope with complex manufacturing tasks. Such robots will eventually be able to tackle multi-part assembly and will have the capacity to adapt to different jobs. Robots capable of moving goods and people will also find wide application in factories, warehouses, hospitals, and within existing transport networks. Simple systems already exist (transit trains, warehousing, dispatch & sorting). Future systems will provide more efficient goods and transport management. This will improve the efficiency of our current transport infrastructure and provide mobility services in hospitals, office blocks, and public places. In all cases, systems will be needed that collect and prioritize requests, dynamically assign routes and missions, manage conflicts and incidents, and monitor robot states and schedule maintenance.

- **Homeland Security & Infrastructure Protection**. Robotics technology offers tremendous potential for applications in border protection, search and rescue, port inspection and security. In addition, robotics technology is expected to be increasingly used to automate the inspection, maintenance, and safeguarding of bridges, highways, water and sewer systems, energy pipelines and facilities, and other critical components of national infrastructure. Robots that undertake these tasks will need to work in every environment. They will require advanced sensing and high-level cognitive capabilities, particularly in regard to their capacity to fuse and manipulate data and interpret objects in their environment. At present, such systems are tasked primarily with the acquisition of information such that this is interpreted by humans remotely. In the longer term, these systems will be expected to identify and respond to unexpected events, isolating potentially dangerous ones and referring these to humans. Increasingly, complex surveillance and security missions will require the deployment and cooperation of multiple robotic systems.

- **Defence**: The scope of potential application for autonomous systems in the Defence domain is quite broad. Examples include tasks in rapid environmental assessment, improvised explosive device detection and defeat, explosive ordnance disposal, countermining, force protection, obstacle clearance, electronic warfare, battle damage assessment, intelligence, surveillance, and reconnaissance, intelligent countermeasures, chemical, biological, radiological and nuclear detection and identification, battlefield simulation and rehearsal, casualty first aid and evacuation, combat search and rescue, logistic support, cargo packaging and pallet assembly, robotic re-arming, convoy duties, re-fuelling, and vehicle recovery.

Given the scope of autonomous systems in the Defence domain, it is worthwhile expanding on some of the more significant applications. To provide additional context, Table 1 and Table 2 provide a breakdown which is categorised against the Defence Operations and Enabling Functions framework.

Table 1: Applications in (relevant) Defence Operations

Applications in Defence Operations	
Cyber Operations	Increasingly intelligent software agents will be applied for the collection, fusion, processing, exploitation and dissemination of intelligence with faster decision cycles. Similarly, agents in the cyber domain will be capable of assuring protection of our networks and systems or alternatively conducting targeted attack.
Control & Denial	Airborne, surface and sub-surface drones could be used for a range of tasks including rapid environmental assessment, hydrographic tasks, mine clearance, search, and persistent surveillance.
Population Centric Operations	Airborne systems will also be able to track bushfires, flood levels, and provide connectivity in disaster impacted or denied environments.

Table 2: Applications in (relevant) Enabling Functions

Application in Enabling Functions	
Command & Control	A wide range of decision support tools will be employed across all ranks. Automation of other tasks will reduce cognitive work-loads and be capable of intelligently undertaking routine tasks. This will reduce operational footprint in headquarter and theatre.
Situational Awareness & Comms	Persistent communications nodes and surveillance systems that can be re-tasked, or adaptively self-task, to improve connectivity and awareness.
Lift & Logistics	Niche applications in tactical airdrop will permit replenishment in inaccessible, remote and dangerous locations without direct risk to personnel. Automation of some aspects of long haul logistics will reduce crewing levels and improve force protection. Refuelling operations could also be automated.
National Support	Virtual software agents will be increasingly employed in training systems and large scale experimentation.

2.4 Legal Considerations[8]

Development of regulatory frameworks for new technologies often lags behind the desire for their employment. For example, the absence of international standards, regulations and procedures to govern the safe integration of UAVs into civil airspace is often cited as key factors limiting growth in the civilian UAV sector (Teal Group, 2011). As a result, most civil operations of UAVs are limited to test or demonstration flights. In general, the novelty of the technology, difficulties with determining the vector of causality and allocating responsibility, and unwillingness to burden the relevant agencies responsible for drafting such laws with additional work combine such that the use of most robots currently falls within a regulatory gap. That is, the technology appears to be under a loose legal framework, is self-regulated, or only allowed to operate in restricted areas.

Autonomous systems will only be integrated into society when they are technically capable of undertaking their defined roles, it makes sound economic sense to do this, and there are organisational and legal frameworks to accommodate them. As technology allows, certain decisions may then be made quicker, more reliably or more consistently by technology and the pressure to remove the human will increase. However, as autonomous systems become more sophisticated issues of liability will become important. Sometimes accidents will happen, but if robots do not demonstrate sufficient and predictable capacity to comply with the legal obligations, questions of liability will result in legal challenge.

[8] This section is cited from Finn (2012, pp. 139-141) with minor changes.

3. Models of Autonomy

3.1 Are Machines Autonomous?

In individuals, autonomy is defined in terms of rationality, determinism, and agency. Much of what we understand by autonomy is then focused around the ability to freely make choices about our actions, motivation and reasoning, and the assumption of responsibility for their outcomes. These ideas do not immediately apply to programmed, robotic systems, however sophisticated or capable. When considering the nature of an autonomous machine, the following questions then become important.

- Are machine capable of reasoned judgement?
- Do machines possess free will?
- How can machines be made accountable for their actions?

The answers to such questions are not immediately apparent or necessarily agreed. For example, autonomy in humans includes the capacity to make an informed decision, without duress or coercion. However, autonomy in machines must surely reflect the notion of fatalism, where the entity is incapable of taking actions other than those it must execute; being predestined or inevitable through programming. For many, questions such as these may seem to be purely philosophical, but notions of self-determinism lie at the very heart of the study of Artificial Intelligence. However, our problem can be explained quite simply. Without tackling these issues, it is simply not possible to determine a meaningful way to assert (positively or negatively) that any system is actually exhibiting autonomy at all.

The idea of autonomy as applied to both individuals and mechanical systems is further compounded by the intricate relationships between properties of the actor, the range of tasks and functions being performed, context and environment, and influence and constraints imposed by all actors. Hence, autonomy in machines cannot be so simply expressed as a property of the system itself, but rather as its contextual on the task at hand, the environment under which it is performed, and the influence and constraints of all the various stakeholders. It is this aggregation of the inter-dependent concepts of capability, its context of use, and relationships to other agents which define the concept.

We begin by asking what it means for an engineered system to be autonomous. As the definition and concept of Autonomy in engineered systems cannot immediately be assumed, or logically follow from, that applying to rational individuals; an approach is employed to develop a set of root clauses based on identifying the underlying principles behind Autonomy. These root clauses form a conceptual basis for the understanding Autonomous Systems and are formed through review of the literature. This is crucial to founding the work in context to other key authors within the field of research, noting their findings and applying them to the particular application of technical forecasting within

Defence and National Security. In doing so, we present a reasoned and informed definition of Autonomy which is fit for our purposes.

Identifying root clauses and principles which apply to Autonomy is additionally useful because it provides a consistent way to assess or evaluate a number of models which have been presented; such as those posed by Sheridan & Verplank (1978) and more recently Cummings (2004). In essence, models which attempt to characterise levels or types of Autonomy in systems often focus on a narrow range of attributes and properties of Autonomous Systems. These models are very useful in some applications but are also limited when applied more broadly. This report looks at the popular models in the field and discusses their range of application. A simple, yet practical model is then proposed which overcomes the restrictions in the published literature.

3.2 Categories of Human-Robot Interaction

Early attempts at developing models for human-machine interaction have predominately focused on the capacity to replace, in full or part, functions which may have previously been carried out by a human. The suggestion here is that categories of automation can be defined in terms of operator burden. These suggest a purely technical scale, which spans the realm of manually operated (remote & tele-operated) systems through to systems which are capable of independent operation.[9] Such a continuum was defined by the US Army Board of Science and Technology (Rose, 2002) and by a Joint DARPA – USAF Project Team (NRC, 2005).

[9] This explicitly excludes systems which might be capable of internalised cognition; that is, the system may be 'fully' autonomous but may not task itself of its own volition.

Table 3: Categories of human-robot interaction

US Board of Army S&T (Rose, 2002) **Focused on mission management, and methods of intervention**	**DARPA-USAF Joint Project Team (NRC, 2005)** **Focused on mission management, and methods of execution**
Level 1 (Manual Operation)	**Level 1 (Remotely Controlled & Tele-operated)**
The human operator directs & controls all functions of the mission.	A human operator controls a robotic vehicle from a distance. The human performs all of the cognitive processes. The on-board sensors and communications enable the operator to visualise the location and movement of the platform within its environment and its on-board effectors enable the human to act on the information it provides.
The vehicle manoeuvers autonomously.	
Level 2 (Management by Consent)	**Level 2 (Semi-autonomous)**
The system recommends courses of action for nominated functions.	These systems have advanced navigation, obstacle avoidance, and data-fusion capabilities to reduce the dependency on and frequency of operator control. They are also designed to adapt, with pre-defined guidelines, to simple changes in mission as designated by an operator.
The system prompts the operator for information as required.	
Today's autonomous vehicles operate at this level.	
Level 3 (Management by Exception)	**Level 3 (Platform-centric Autonomous)**
The system automatically executes mission-related functions when response times are too short for operator intervention.	A fully autonomous platform can undertake complex tasks, and identify and request the information required to complete those tasks. It can also respond to and perform mission planning against new commands issued by an operator.
The operator may override or redirect execution of actions at will.	
Exceptions are brought to the operator's attention for intervention.	
Level 4 (Fully Autonomous)	**Level 4 (Network-centric Autonomous)**
The system automatically executes mission-related functions when response times are too short for operator intervention.	These systems have sufficient autonomy to operate as independent nodes within the context of a network-enabled force. They are capable of communicating with the network, incorporating the relevant information it in their mission planning and execution, and responding to other information requests, including the resolution of conflicting commands.
The operator is alerted to function progress.	

3.3 Communication as a Measure of Autonomy

Categories of human-robot interaction are an easy, although perhaps simplistic, way to differentiate between levels of autonomy. Models like this are widely cited in the literature. However, the inherent limitations to these model are that they do not scale well with complex systems; having many operators, many machines, or both. For example, a current growth area in unmanned systems technology is that of 'swarming' systems. This already brings some difficulties to the earlier models of human-interaction in that a swarming system has inter-platform connectivity. This means that a further scale of autonomy, being that of machine-to-machine collaboration, is needed. Cummings (2004) is renowned for modelling autonomy across networks and machines. This model is presented in Table 4.

Table 4: Categories of machine communication

Cummings (2004) **Focused on networks, and machine communication**		
Maximum Network Autonomy	↑	Vehicles are in full collaborative communication Individual tasking changes according to a predetermined algorithm There is no human intervention
		Vehicles collaborate with one another and human operators Human operator interacts only with 'lead' vehicle(s)
		Vehicles communicate with one another for de-confliction Vehicles dependant on human for new mission tasking
Minimum Network Autonomy	↓	Vehicles do not communicate with one another Vehicles follow original tasking until human intervention

Cummings work can be interpreted as a significant step forward in the understanding of how autonomy might be measured. It does not replace earlier work but provides a second and complementary perspective. Both a machine-centric and an operator-centric approach can be employed to provide greater balance when assessing or measuring autonomy.

3.4 Degrees of Automation

The work of Sheridan and Verplank (1978) is perhaps the most well-known of all models of automation. They introduced a scale for NASA which defined ten levels, which is considered by many to be a seminal piece. The work serves as a reference benchmark and is still relevant today.

Table 5: Levels of automation based on decision-making

Automation Level		Sheridan & Verplank (1978) Focused on outputs, and selection of options	Endsley & Kaber (1999) Focused on inputs, and generation of options
1.	**Manual Control**	The computer offers no assistance; human must do it all.	The human monitors, generates options, selects options (makes decisions) and physically carries out options.
2.	**Action Support**	The computer offers a complete set of action alternatives, and…	The automation assists human with execution of selected action, while the human still performs some control actions.
3.	**Batch Processing**	…narrows the selection down to a few, or…	The human generates and selects options which are turned over to automation to be carried out.
4.	**Shared Control**	…suggests one, and…	Both the human and the automation generate possible decision options but the human has control of selecting which options to implement (executing options is a shared task).
5.	**Decision Support**	…executes that suggestion if the human approves, or…	The automation generates decision options that the human can select and once selected, the automation implements it.
6.	**Blended Decision**	…allows the human a restricted time to veto before automatic execution, or…	The automation generates an option, selects it and executes it if they human consents (the human may approve of the option, select an alternative or generate another).
7.	**Rigid System**	…executes automatically, then necessarily informs the human, or…	The automation provides a set of options, the human has to select one of them, and once selected the automation carries out the function.
8.	**Automated Decision**	…informs him after execution only if he asks, or…	The automation selects and carries out an option but the human can supply input as options are generated.
9.	**Supervisory Control**	…informs him after execution if it, the computer, decides to.	The automation generates options, selects and carries out a desired option; while the human monitors the system and intervenes if needed.
10.	**Full Automation**	The computer acts autonomously, ignoring the human.	The system carries out all actions.

Sheridan and Verplank's model focuses not on either the operator or the machine but the authority and responsibility for decision-making. This type of approach is well received in Defence as it naturally reflects common ideals such as supporting personnel through technology. This model emphasises the selection of options, typically relating to mission planning, and the automated execution of that plan. In their model, a system is autonomous if it has the capacity to filter, prioritise or select between alternative courses of action.

Endsley and Kaber (1999) added further refinement to Sheridan and Verplank's model. They also propose ten categories but their emphasis is placed on the generation of options. This is important because it distinguishes between who conducts mission planning activities, be it human or machine, and the function of selecting actions. Thus, Sheridan and Verplank's model has greatest relevance where measurement of output or outcomes is important. Likewise, Endsley and Kaber's model has greatest relevance where measurement of input or formulation is important.

3.5 Task Models of Automation

Table 6: Assessment scale based on stages of information processing

Parasuraman et al (2000) Focuses on information, and states of processing	
A.	**Information Acquisition** Automation of information acquisition applies to the sensing and registration of input data. These operations are equivalent to the first human information processing stage, supporting human sensory processes.
B.	**Information Analysis** Automation of information analysis involves cognitive functions such as working memory and inferential processes. At a low level, algorithms can be applied to incoming data to allow for their extrapolation over time, or prediction.
C.	**Decision Selection** Decision and action selection, involves selection from among decision alternatives. Automation of this stage involves varying levels of augmentation or replacement of human selection of decision options with machine decision making.
D.	**Action Implementation** Action implementation refers to the actual execution of the action choice. Automation of this stage involves different levels of machine execution of the choice of action, and typically replaces the hand or voice of the human. Different levels of action automation may be defined by the relative amount of manual versus automatic activity in executing the response.

Previous models explored autonomy through the lenses of operator-interaction, machine-communication, and decision-making. One final refinement is offered by Parasuraman et al (2000); that is, to define automation in terms of the accomplishment of tasks.

Parasuraman et al is unique in that no specific scale is defined. Non-categorical ranking across a number line is possible and perhaps even preferred. This can make the model difficult to apply due to subjective differences of interpretation. The approach is also unique in one other aspect. Previous models have struggled to capture how autonomy changes with the mission or task being undertaken. Parasuraman et al use four stages of information processing as their motivating framework and rightly observes that autonomy is not fixed but changes based on the context of assessment.

A recent revision of this approach was presented by Proud et al (2003). Proud et al adopt Boyd's 'OODA' loop as the four motivating tasks. This is also based on information processing and widely recognised within the Defence community. This framework extends Parasuraman et al by providing eight categories, per task, against which autonomy can be compared. It is presented in Table 7 and Table 8.

Table 7: Assessment scale based on OODA loop (Levels 1-4)

	Proud et al (2003) Focus on role of human and machine across stages in the OODA loop			
	Observe	**Orient**	**Decide**	**Act**
1.	Human is the only source for gathering and monitoring (defined as filtering, prioritizing and understanding) all data.	Human is responsible for analysing all data, making predictions, and interpretation of the data.	The computer does not assist in or perform ranking tasks. The human must do it all.	Human alone can execute decision.
2.	Human is the prime source for gathering and monitoring all data, with computer shadow for emergencies.	Human is the prime source of analysis and predictions, with computer shadow for contingencies. The human is responsible for interpretation of the data.	The human performs all ranking tasks, but the computer can be used as a tool for assistance.	Human is the prime source of execution, with computer shadow for contingencies.
3.	The computer is responsible for gathering and displaying unfiltered, un-prioritised information for the human. The human still is the prime monitor for all information.	Computer is the prime source of analysis and predictions, with human shadow for contingencies. The human is responsible for interpretation of the data.	Both human and computer perform ranking tasks, the results from the human are considered prime.	Computer executes decision after human approval. Human shadows for contingencies.
4.	The computer is responsible for gathering the information for the human and for displaying all information, but it highlights the non-prioritized, relevant information for the user.	The computer analyses the data and makes predictions, though the human is responsible for interpretation of the data.	Both human and computer perform ranking tasks, the results from the computer are considered prime.	Computer allows the human a pre-programmed restricted time to veto before execution. Human shadows for contingencies.

Table 8: Assessment scale based on OODA loop (Levels 5-8)

	Proud et al (2003) **Focus on autonomy assessed across the OODA loop**			
	Observe	**Orient**	**Decide**	**Act**
5.	The computer is responsible for gathering the information for the human, but it only displays non-prioritized, filtered information.	The computer overlays predictions with analysis and interprets the data. The human shadows the interpretation for contingencies.	The computer performs ranking tasks. All results, including 'why' decisions were made, are displayed to the human.	Computer allows the human a context-dependant restricted time to veto before execution. Human shadows for contingencies.
6.	The computer gathers, filters, and prioritizes information displayed to the human.	The computer overlays predictions with analysis and interprets the data. The human is shown all results.	The computer performs ranking tasks and displays a reduced set of ranked options while displaying 'why' decisions were made to the human.	Computer executes automatically, informs the human, and allows for override ability after execution. Human is shadow for contingencies.
7.	The computer gathers, filters, and prioritizes data without displaying any information to the human. Though, a 'program functioning' flag is displayed.	The computer analyses, predicts, interprets, and integrates data into a result which is only displayed to the human if result fits programmed context (context dependant summaries).	The computer performs ranking tasks. The computer performs final ranking and displays a reduced set of ranked options without displaying "why" decisions were made to the human.	Computer executes automatically and only informs the human if required by context. It allows for override ability after execution. Human is shadow for contingencies.
8.	The computer gathers, filters, and prioritizes data without displaying any information to the human.	The computer predicts, interprets, and integrates data into a result which is not displayed to the human.	The computer performs ranking tasks. The computer performs final ranking, but does not display results to the human.	Computer executes automatically and does not allow any human interaction.

4. Developing a Foresight Framework

4.1 Extending Existing Models of Automation

Models of interaction, such as those presented in Section 3, focus on the accomplishment of a task and measuring the degree to which the autonomous systems rely on external control. This is both a reasonable representation and practical, for this facet of human-machine interaction. However, humans also interact with systems in other ways not relating to information or control. Other facets of the human-machine interaction need to be considered. In particular, the degree to which an autonomous system requires physical, as opposed to informational, support is important. Systems which require manual launch and recovery, extensive maintenance, and refuelling cannot be said to be as autonomous as those that do not. This is often overlooked since these tasks are rarely considered to be part of completing the mission.

A framework of three axes can be constructed to include a support axis along with the Sheridan-Verplank and Cummings et al models. This would comprise of three dimensions to describe the degree of autonomy a machine has from human control, machine control and physical support. This approach encompasses the cost of all physical inputs. Across the life cycle of the system; the logistics of spare parts, warehousing and disposal can be measured as aspects of autonomy which as just as valid as those based on the service exchange of information. However, while using one or three scales for degrees of autonomy is convenient and useful for certain purposes, it is still insufficient to use such scales to completely characterize the relationship between the machine and other entities that define its autonomy. In truth, there are more than just the three dimensions to autonomy and measurement of autonomy changes with task and context. [10]

4.2 Task and Environment

Entities do not exist in isolation. Rather, they co-exist in an environment and share (or compete) for resources. The actions of one entity alter the state of world, which directly or indirectly affects other entities. Further, the issue of consent, rules and regulations is critical to understanding autonomy, particularly when humans are involved. Autonomous systems are artificial, designed and created by humans to satisfy human goals. Once assigned goals, they execute a number of steps in sequence. In simple terms, they determine if the current state meets the current goals, then they determine the appropriate actions necessary to reach those goals, and finally they execute those actions. We define an *action* to be something that an entity does to change the state of itself, other entities or its environment. By this definition, actions are always deliberate even if they are, at times, spontaneous. A *state* is a particular condition that entities and/or the environment are in at some particular time. A *task* is an action that is taken by one more entities, to achieve some *goal* (desired state).[11]

---

[10] This idea is described in greater detail within Finn & Mekdeci (2014, p.30).

[11] The notions of 'state', 'action', 'objective', and 'state' are discussed in Finn & Mekdeci (2014, p.15). Their work discusses both task and objective hierarchies. In addition the discussion of the 'context'

* **Physical Environment.** Real systems (those that are not virtual and exist only in software or as a concept) reside in, and operate within a physical environment. If a system is part of a larger system, or is nested within a System of Systems, then its environment may or may not be conducive to the system's operations. Terrain, weather and obstacles often influence a machine's autonomy. While an unmanned system may be able to autonomy's navigate a particular terrain in perfect weather, the reduced visibility of a snowstorm may necessitate additional localization information from a human operator, reducing the machines autonomy as a result.

* **Exogenous Entities**. Systems often do not reside alone within their environments, and typically interact with other entities intentionally or unintentionally. Other entities may include humans, animals or other machines and these entities may be friendly, neutral or hostile. The existence and behaviour of exogenous entities may greatly impact a machine's autonomy. For example, if there are multiple victims to rescue, then a search & rescue robot may not have the ability or the authority to prioritize which victims should be rescued first, and may depend on human operators for that particular decision.

* **Stakeholder Constraints.** Stakeholders may impose constraints that influence a machine's autonomy. For example, when life-and-death decisions are being made, regulations may enforce supervisory control by a human operator, regardless of the machine's capability. As trust and trustworthiness progresses to the point when stakeholders are more confident in the machine's capability to make life-and-death decisions, the regulations may be the opposite and the machine will make the decision autonomously, regardless of human input.

Efficacy is the system's physical ability to perform an action, such as sense the environment, move between two points, or manipulate an object. Efficacy can be reflected by not only whether the entity can actually perform the task, but how well it can perform it (Finn & Mekdeci, 2014, p.15). On a low level, efficacy can be measured, for example, by how much weight the machine can lift, how fast an unmanned plane can fly, or how many client requests a web server can process per unit time. For higher level tasks, such as performing maritime security, efficacy can be more difficult to measure but can be approximated with metrics such as the percentage of arriving boats detected, average time to identify target, and false alarm rate.

Efficacy for most tasks implies that there is some form of intelligence, either by the entity itself or by another entity controlling it. This is because actions that an entity take belong to a hierarchy of tasks, with numerous lower level actions required to complete higher level tasks. For example, a rescue-mission involves several lower level tasks, such as moving from point A to point B and detecting victims within a particular environment. The task of recovering victims might be a sub-action of a higher-level task, such as performing maritime security. Every task can be broken down into a set of sub-tasks, each

---

includes the Physical Environment, Exogenous Entities and Stakeholder Constraints. These have been cited from pages 45 and 46 of that same reference.

one at a lower level than the rest, until all that is left are only primitive functions that are performed by components and not sub-systems. Therefore, even if an entity was instructed to perform task A and lacked the intelligence to decide that A was the task that should be performed, the entity may still require some degree of intelligence to execute the sub-tasks required to complete task A.[12]

4.3 Trust in Autonomy[13]

In measuring autonomy, accounting for the capability of the entire system important. Humans are an integral part of this system and the competency, attitude and behaviour of the human should be reflected. Merritt and Ilgen (2008) demonstrated how variations in trust, between operators of autonomous systems, impact on perceived measurement of that systems autonomy. The conclusion being that, increasing the capability of the automation is not sufficient to ensure effective use. Instead, it is necessary to increase trust between humans and autonomous systems to ensure that features of automation are appropriately employed.

Trustworthiness is how well a machine performs a particular task without human intervention, i.e. the capability of the automation. *Trust* is how well humans perceive the automation is performing a task (or vice versa). Ideally a human should trust a machine as much as it is trustworthy, no more or less. Lee and See (2004) refers to calibration as how similar a human's trust of the automation is to its trustworthiness, and *calibrated trust* as the ideal situation where an operator trusts the automation just as much as is warranted.

In many cases, calibrated trust is poorly aligned to the system and its performance is suboptimal as a result. This occurs in two ways. First, *disuse* occurs when an operator trusts a machine less than it is actually trustworthy, leading to situations where the automation is underused. Second, when an operator trusts the automation more than it reasonable, then *misuse* occurs. Hence, trust in autonomy is a multi-dimensional construct that changes with time. It is influenced by the expression of purpose, intention and role; approaches to developing and determining trust; functional aspects such as system capability; and reliability within an operating environment.

4.4 A Revised Model for Technology Foresight

To assess the impact of technology changes on machine autonomy, one must take a look at the factors that play a role in determining the human interaction required to perform a particular task are:

1. the systems base capabilities;
2. the task itself;
3. the context under which the system performs that task; and

[12] Cited from Finn & Mekdeci (2014, p.19) with minor changes.
[13] Trust in autonomy is discussed in detail in Finn & Mekdeci (2014, pp. 74-86).

4. the trust between parties in the system.

Advances in technology are likely to have a direct impact on the capability of the system and perceptions of trustworthiness. In some cases, they may also provide the ability to perform new tasks. Thus, by introducing effective novel and new technology the capability of the system is increased and:

- the autonomy of the machine can be directly increased; and/or
- the complexity of tasks that the system can performs can increase; and/or
- the system can operate in environments which are more complex; and/or
- stakeholders will have trust in the system.

This illustrates the interdependence between advances in technology and what we know of autonomous systems and their performance. It is also possible to observe, as a corollary, that if a task or contextual complexity is increased that autonomy will decrease and/or trust in the system will be diminished.

Four key attributes are defined; system, mission, context and trust. Each of these attributes is comprised of a number of factors which should be considered when assessing the autonomy of a system. Studies in autonomy from the University of South Australia (Finn, 2011; Finn, 2012; Finn, 2014; and Finn & Mekdeci, 2014) have derived those factors from their roots in the literature. In partnership, that program of research has been synergised into a single coherent taxonomy. Table 9 presents the proposal for a revised autonomous systems taxonomy, or conceptual model, which presents an alternative to those in Section 3. This can be used to assess the autonomy of a system at a holistic level. It promotes greater understanding of the characteristics which contribute to making a system autonomous and is perhaps more amenable to practical application as a result.

- **System.** The factors under the system component are derived from Finn (2011) as the outcomes of a symposium of experts at the 2011 Defence Effects conference.[14] The autonomous systems workshop aimed to "identify the drivers and functional requirements that permit development of autonomous systems with the capacity to operate reliably and safely in dynamic environments shared by people, property, and other systems". The outcomes additionally informed the projections in Appendix C.

- **Mission and Context**. The factors under mission and context were largely explored in Finn (2012) in the report *Implications for Autonomy & Autonomous Systems*. The objective of this report was to "examine the current and likely future systems challenges and technological impediments that may prevent achievement of robust autonomy capable of integration into a human-centric future society." An

[14] Edinburgh Parks, Adelaide, Australia, 9 November 2011.

additional workshop was held in 2014[15] to specifically target capabilities for the ADO (Finn, 2014).

- **Trust**. Trust and human-machine interactions was the primary focus of Finn & Mekdeci (2014). This work also informed the conceptual model directly, being a "list of system, task and contextual factors that define how autonomous a system will be when it performs a task in a particular context." This focuses on understanding what makes systems trustworthy and articulating this in a list of contributing components.

[15] Mawson Lakes, Adelaide, Australia, 4 February 2014.

Table 9: Proposed new taxonomy for autonomous systems

Component	1. System	2. Mission	3. Context	4. Trust
Constituent Elements	Actuation & Locomotion Architecture & Integration Computation & Processing Grasping & Manipulation Mapping & Navigation Power Management & Energy Sensing & Perception Signature / System Footprint Technical Survivability	Frequency of Action Required Precision Synchronisation of Effects Task Complexity Time Constraint	Adversarial Intervention Complex Structures EM or RF Interference Environment & Change Hazards Policy & Regulation Stakeholders Visibility, Light & Obscuration	Certification Cooperation & Collaboration Competency & Training Human & Robot Interaction Learning & Adaptation Past Experiences Perception Predictability Responsibility Task Complexity Usability
Significant Trends	Component Miniaturisation Cyber & EM Hardening Open Architectures Reduced Cost	Asymmetric Operations Faster Tempo Media & Public Relations Operations Other than War	Concern over CBRN Discrimination Problems Legislative Issues Need for Sophistication	Commercial Drivers Early Adversarial Adoption Investment Increasing Proliferation of Technology
Critical Enablers	*Research Expertise* — *Developments in science, convergence of research, transformational advances* *Industry & Government Support* — *Economic benefit, novel applications, integration, commercialisation* *Defence Strategy* — *Concepts of operation, legal & ethical commitments*			

5. Applying the Conceptual Model

5.1 Technology Foresight

Technology Foresight is the field of scientific regard which investigates the emergence, performance, or impact of technology across society. It aims to describe usage or uptake, and evolving trends, in technological development over time. Through doing so, it seeks to determine the implications of those developments both current and forthcoming. Research outcomes of technology foresight then exist in a continuum from the present into the future. As such, the field is never static, constant evolving, with its different forms and potentialities connected through time.

Australia capitalises on the opportunities presented by technology foresight, in building national power through investment and maintaining its regional capability edge. Likewise, technology foresight identifies opportunities to mitigate against the risks inherent in strategic shock and technology surprise. Understanding emerging scientific and technological trends, within the wider socio-economic and strategic environment, and their implications for Defence and National Security, becomes critical. This aspect of necessary and measured response to change is then one of the key drivers to guide new policy, strategy, future concepts, and force development.

As a key driver to strategic decision-making processes, the rational basis supporting technology foresight itself becomes important. Confidence in decision-making is underpinned by confidence in the argument upon which it is enacted. In this sense, technology foresight must be capable of providing robust and auditable outcomes. Ironically, the strength of technology foresight cannot be measured until after the fact. By the very nature of prediction, technology foresight analyses causal chains of events which have yet to be observed. It is argued that strength of argument is then the primary means by which technology foresight is measured.

> *Since the purpose of a technology forecast is to aid in decision making, a forecast may be valuable simply if it leads to a more informed and, possibly, better decision. A forecast could lead to decisions that reduce future surprise, but it could also inspire an organization to make decisions that have better outcomes—for instance, to optimize its investment strategy, to pursue a specific line of research, or to change policies to better prepare for the future.*
>
> (NAP, 2010)

Vanston (2003) concludes that if the process of technology foresight leads to a decision which is better, more informed, or rationally argued, then it has been successful in achieving its intent. For our purposes, we wish to advise decision makers about autonomous systems technologies, so that they have a reasoned understanding of their capability and can make informed decisions about their potential applications and operational impact. For this we need to apply our conceptual model.

5.2 Methodological Approach

In order to appreciate the application of our conceptual model for technology foresight, it is worthwhile revisiting the methodological approach. This is comprised of six stages.

1. **Trends in autonomy research reports**.

   We started by investigating the factors which contribute to system autonomy and understanding how technologies might contribute to those factors. In doing so, we can monitor developments in the technologies or pre-propose milestones which must be logically achieved in the evolution of autonomous systems. An extensive program of research was conducted in partnership with the University of South Australia. This resulted in three foundation publications (Finn, 2011; Finn, 2012; and Finn & Mekdeci, 2014) and reported major themes in autonomy.

2. **Derivation of the root clauses and key factors of autonomy**.

   The next step in this approach is to identify factors of autonomy from major themes, together with any related concepts, and the contexts in which they apply. With such an understanding, those factors can then be investigated through an experimental process, to examine varying hypotheses about how they relate or contribute to capability. This was accomplished through derivation of the root clauses and key factors of autonomy was performed through extensive literature review of existing models, as per Section 3 of this report, and recommendations from the University of South Australia in step 1.

3. **Decomposition of those factors into a hierarchy**.

   Section 4 of this report provides the resulting conceptual model to be used for technology foresight. It is based on four capstone factors of the system capability, mission complexity, context of employment, and extent of operator trust in the system. Each of the four capstone factors is broken into its most significant constituent elements and a summary statement of current observable trends. This is provided in Table 9 of this report.

4. **Definition of a parametric model of investigation**.

   In this parametric model, changes in named factors of specific interest can be studied. This also forms the principle mechanism by which the model can be tested and refined. It can be used in an elicitation process, workshop, or experimental campaign for knowledge capture and representation. This includes key perspectives from discussion, to categorise statements and options into logical groupings, and provide consistency of approach. The model is amenable to additional analysis because it provides a way to identify gaps in knowledge, where additional scrutiny or information is required. Through application, the model itself can be refined, changed or improved as a result. This would tell us more about, and enrich or understanding of, what makes systems autonomous.

5. **Case studies using the parametric models**.

   This is the first point in the methodology in which the model is employed for the purposes of technology foresight. It is suggested that a set of selected autonomous systems is mapped onto the model. In mapping these capabilities across the model, the specific technologies which contributing to each factor in the model can be examined and projected into the future based on reported trends. Through this mechanism it is possible to consider 'vectors' across the space to show how technological innovation would impact on the factors in autonomy.

6. **Statement of impact**.

   Application of the model for technology foresight allows us to identify technologies with potential for significant impact in operations. This provides indicators-and-warnings for Australia and its allies about risks and opportunities in technologies for autonomous systems. It supports informed decisions about investment, acquisition, and research in those technology areas.

Of these six phases, the fourth stage largely describes a stage of test and evaluation. In this, the model itself is subjected to examination and refined. The last two describe the application of the conceptual model for technology foresight. Stage five is the foresight activity and stage six reflects the action of reporting and informing significant stakeholders. This naturally leads to the following three recommendations for further work, to take the conceptual model and applying it.

5.3 Recommendations

5.3.1 Recommendation 1: Adequacy of Framework

Application of the Autonomy framework itself serves to test the base conceptual model; that is, to determine its adequacy for the task. When employed, it may be that elements of the framework have been missed or undue emphasis has been placed on some elements which are not needed. Application then assists to expand and deepen our knowledge of Autonomy as well as to improve the conceptual model.

Recommendation	
1.	The Autonomous Systems Framework should be tested, for its fitness for use, within the context of an evaluation workshop.

5.3.2 Recommendation 2: Parametric Study

A complete model also naturally lends itself to parametric investigation, against which existing and future Autonomous Systems might be mapped. This assists us in the process of technology forecasting because it might be possible to make reasoned deductions (or at least plausibly infer) how technology might contribute or impact on the factors expressed

in the conceptual model. If so, the model can be used to provide the types of indicators and warnings for changes in Autonomy.

Recommendation	
2.	Conceptual future systems be generated through parametric variation of elements in the framework. Explore the range of systems that emerge bad make estimates for how, when or if such systems will be operationalised.

5.3.3 Recommendation 3: Understanding Operational Impact

There will always be development paths that could not be predicted through the application of our framework, and we call this technology surprise.[16] However, our framework is useful in exploring possibilities, to the extent that is possible. The question of how the ADO positions itself in preparation for technology surprise and also during an unforeseen event is the key issue. Applying our framework can help the ADO manage potential sensitivity, exposure and risk. Technologies with greatest potential impact on Autonomy can be then investigated. This includes the capability which might be developed if those technologies were realised (by the ADO or by an adversary).

Recommendation	
3.	Outcomes of Recommendation 2 be investigated for their potential utility to the ADO.

[16] This can occur through unforeseen technological breakthrough, disclosure of an advanced clandestine development program, otherwise unanticipated and accelerated pace of development, and novel or innovative application of existing technology.

6. Conclusion

Technologies for autonomous systems are important to Defence because they have the potential to extend the reach and capability of traditional forces while reducing operational footprint and threat to personnel. This increases access to regions of operational interest, especially in contested regions, possibly enhancing control, freedom of manoeuvre, and denying the area to adversary forces. Measurement of the extent to which a system is autonomous is then of significant interest. If autonomy can be increased then a greater benefit or range of effects might potentially be derived.

However, previous attempts at describing or quantifying autonomy within the literature have been limited in their scope of regard. The true value of autonomy lies within an appreciation of its context and, in and of itself, it is not a solution to any problem. Any model of autonomy must then be sufficiently rich as to capture all aspects of the capability. This must include an appreciation of four key limiting attributes.

1. First, the system has to have the basic capability for autonomy. This means it must have the necessary components be they actuators, sensors, software, and the like that allow it to operate.

2. Second, the task itself must be within the capacity of the system to fulfil. Highly complex tasks necessarily require an equally sophisticated system. If the sophistication of the system is not sufficient then it will increasingly require operator intervention.

3. The context in which the system is employed must also be conducive. A system operating autonomously in one environment may not necessarily be capable of operating autonomously in another.

4. Finally, trust in automation is a multi-dimensional construct, with organizational, sociological, psychological and neurological influences all playing a role. Operators should not trust a system more than it is trustworthy, nor should they trust it less.

Each of these attributes is comprised of a number of factors which should be considered when assessing the autonomy of a system. Working in partnership with the University of South Australia, DST Group has developed a single coherent taxonomy of these factors which is presented as a conceptual model in Table 9.

The purpose of this conceptual model is to support the ADO technology foresight process. This model provides a systematic and auditable way to explore emerging technologies for autonomous system by means of parametric investigation. By exploring each of the contributing factors, it is possible to identify technologies which affect those factors, and to theorise about their impact on operational capability.

The outcomes of this technology foresight activity raises opportunities to mitigate against the risks inherent in strategic shock and technology surprise, build national power

through smart investment, and maintain its regional capability edge through early acquisition. It then informs decision makers in Defence and National Security to shape the future of policy, strategy, emerging concepts, and force development in autonomous systems and their related technologies.

7. References

Boey S, Hugh B, Tailby D and Wheeler S (Eds.) (2014a). *Strategic Science and Technology Estimate 2014*. DSTO, Joint and Operations Analysis Division, Classified Report.

Boey S, Dortmans P, Nicholson J (2014a). *Forward 2035 – DSTO Foresight Study*. DSTO, Joint and Operations Analysis Division, Commonwealth of Australia, Defence Science Communications, Canberra, ACT.

Cebon P and Samson D (2008). *Technology Assessment: Methods and Approaches for Military Applications.* Report developed in partnership with DSTO. University of Melbourne, Parkville, Australia.

Cummings ML (2004). Human Supervisory Control of Swarming Networks. *Proceedings of the 2nd Annual Swarming Autonomous Intelligent Networked Systems Conference*, pp. 1-9.

Defence Science and Technology Organisation (2014). *Defence Science Partnering Deed*. Agreement between the Commonwealth of Australia and the University of Melbourne, 7 October 2014.

Defence Science and Technology Organisation (2015). *Use of Unmanned Platforms by the ADF*. Submission to the Senate Foreign Affairs, Defence and Trade References Committee Inquiry, February 2015.

Ellis M, Palfreyman A, Keir R, Hayward J, Winter R, Dean K (2005). Military Implications of Future Autonomous Systems. *Proceedings of the Land Warfare Conference*, Brisbane, Australia, October 2005.

European Robotics Technology Platform (2009). *Robotic Visions: To 2020 and Beyond. The Strategic Research Agenda for Robotics in Europe*, EUROP, Brussels, Belgium.

Finn A (2011). *Implications for Autonomy & Autonomous Systems: Report on the Autonomous Systems Workshop from the Defence Futures and Emerging Capability / Technology Symposium 2011*. Report developed in partnership with DSTO. Document ID: DA-SC-080-D0064. University of South Australia, Mawson Lakes, Australia.

Finn A (2012). *Implications for Autonomy & Autonomous Systems: Trends in Autonomous Systems.* Report developed in partnership with DSTO. Document ID: DA-SC-080-D0069. University of South Australia, Mawson Lakes, Australia.

Finn A (2014). *Report on Trusted Autonomy Workshop: Tuesday, 4th February 2014*. Report developed in partnership with DSTO. Document ID: DASI-2014-CR-005. University of South Australia, Mawson Lakes, Australia.

Finn A and Mekdeci B (2014). *Report on Trusted Autonomy.* Report developed in partnership with DSTO. Document ID: BM2014. University of South Australia, Mawson Lakes, Australia.

Finn A and Scheding S (2010). *Challenges for Autonomous & Unmanned Vehicles*. Springer.

Huang H-M (ed.) (2008). *Autonomy Levels for Unmanned Systems (ALFUS) Framework. Volume I: Terminology.* Version 2.0. National Institute of Standards and Technology, Gaithersburg, MD.

Lee JD and See KA (2004). Trust in Automation: Designing for Appropriate Reliance, *Human Factors*, Vol. 46(1), pp. 50-80.

Lomuscio AR (2015). *Trusted Autonomous Systems*. Department of Computing, Imperial College London, Research Project (Abstract), Reference EP/I00520X/1.

Mass Technology Leadership Council (2009). *The Massachusetts Robotics Revolution: Inspiring Innovation, Driving Growth and Competitiveness in Leading Industries.* The Innovation Institute, Burlington, MA.

McFarland T (2015). How Should Lawyers Think About Weapon Autonomy?. Accepted for publication in the forthcoming quarterly journal of the *International Review of the Red Cross*.

Merritt SM and Ilgen DR (2008). Not all trust is created equal: Dispositional and history-based trust in human-automation interactions. *Human Factors,* Vol. 50(2), pp. 194-210.

National Research Council (2005). *Autonomous Vehicles in Support of Naval Operations*. Committee on Autonomous Vehicles in Support of Naval Operations, Naval Studies Board, National Research Council of the National Academies, National Academies Press, Washington, DC.

Rose MF (chair) (2002). *Technology Development for Army Unmanned Ground Vehicles*. Committee on Army Unmanned Ground Vehicle Technology, Board of Army Science & Technology, National Research Council of the National Academies, National Academies Press, Washington, DC.

Scofield TJ (2006). *Manning and Automation Model for Naval Ship Analysis and Optimization*. Masters Dissertation, Department of Aerospace and Ocean Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Sheridan TB and Verplank W (1978). *Human & Computer Control of Undersea Teleoperators*. Man-Machine Systems Laboratory, MIT, Cambridge, MA.

Spinks GM and Alici G (2008). *Electroactive Polymers for Micro-Systems: Applications for Actuating Micro-Systems.* Report developed in partnership with DSTO. University of Wollongong, Wollongong, Australia.

Teal Group (2011). *World UAV Market Profile and Forecast*. Teal Group Corporation, Fairfax, VA.

Trevelyan JP, Kang S and Hamel WR (2007). *Robotics in Hazardous Applications.* Report developed in partnership with DSTO. University of Western Australia, Crawley, Australia.

Trevelyan JP (2007). *Robotics Technology Forecast. Mines, Improvised Explosive Devices and Countermeasures*. Report developed in partnership with DSTO. University of Western Australia, Crawley, Australia.

Tweedale J, Bender A and Palmer JL (2013). Proceedings of DSTO's 2013 Workshop on Automation, Autonomy and Autonomous Systems (AAAS). DSTO, Joint and Operations Analysis Division, Limited Distribution Report, DSTO-GD-0811.

Wheeler, S (2015). *Trusted Autonomy*. Emerging and Disruptive Technologies Brief. An initiative of the Technologies Forecasting and Futures Group, Defence Science and Technology Organisation, Fairbairn, ACT, Australia.

Zimmer UR (2007). *Robotics Scoping Report 2007*. Report developed in partnership with DSTO. Australian National University, Acton, Australia.

Appendix A: Autonomy Program

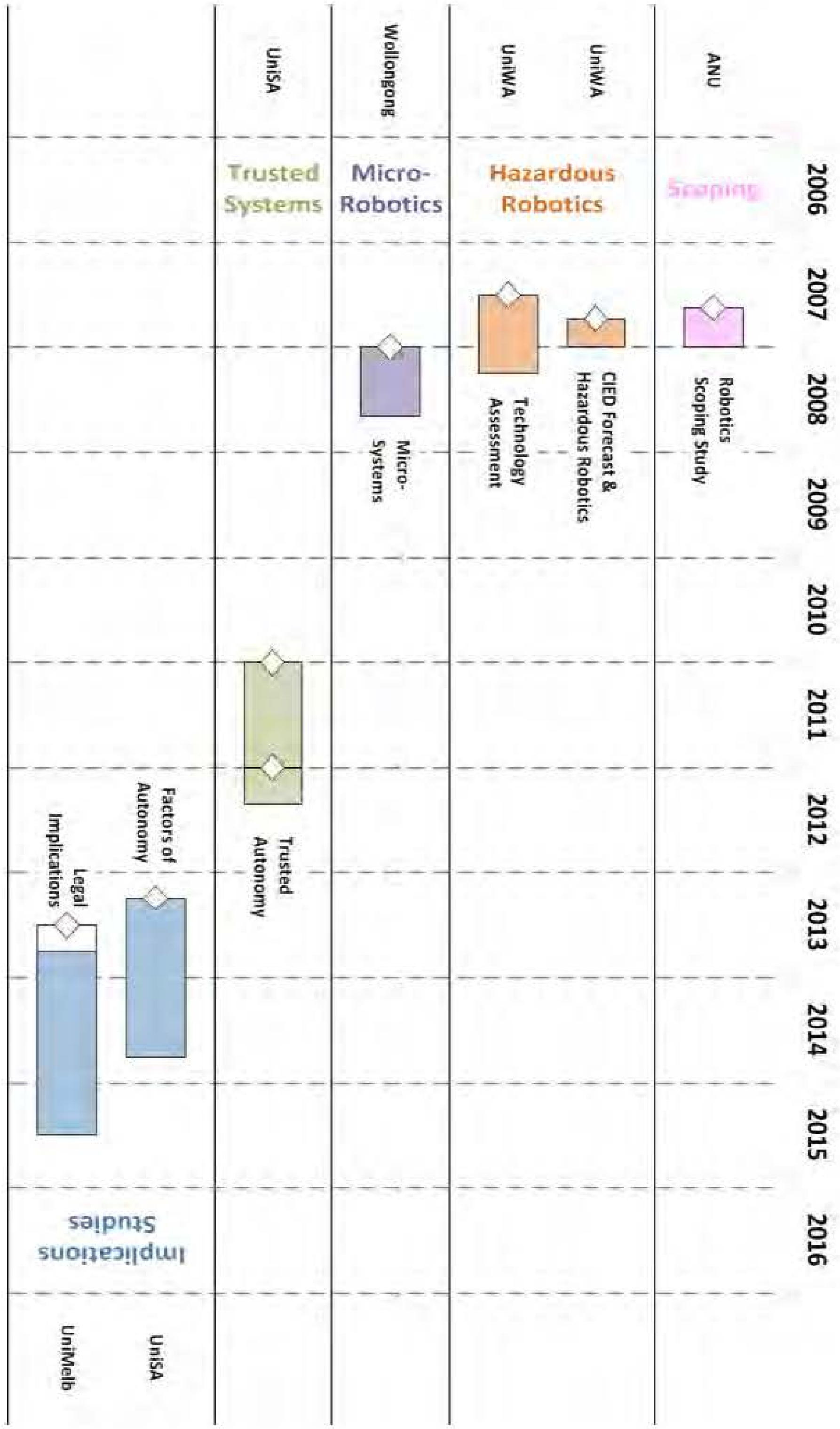

Figure 1: Autonomy program (diamonds mark contract initiation date)

Appendix B: Developmental Timeline[17]

SHORT TERM	MEDIUM TERM	LONG TERM
NAVIGATION		
Safe navigation in unstructured 2-D indoor/outdoor environments Robust probabilistic location techniques (particle filters, grid based map) Avoidance of slow-moving and collaborative dynamic obstacles Navigation based on multi-sensor fusion (all environments) Learn semantic maps through exploring & human input Safety aspects of navigation considered	Greater consideration given to safety aspects for faster motion Understand environment from perception, interaction & human instruction Localisation based on perception of general environmental features Vision-based approaches gain importance using low cost sensing Development of techniques capable of exploiting low cost sensing 2.5-D mapping becomes efficient and accurate for large scale (few km) sectors 3-D mapping is efficient and accurate for small scale (100m) sectors Motion planning techniques for large scale maps and dynamic environments Cognitive approaches to motion planning and re-planning Obstacle avoidance integrated with regulatory frameworks Capacity to navigate in populated areas (e.g. malls, corridors, etc.) Capacity to consider social rules, standards, common human reactions	Safe and reliable high speed, collision-free 3-D (outdoor) navigation Certification procedures developed (e.g. airworthiness & airspace integration) Capacity to operate in novel, open, unstructured and dynamic environments Low cost, safe navigation based on exteroceptive sensing approaches Vehicles able to operate in multiple environments Object representations with symbolic representation Respond to dynamic changes in environment in a manner consistent with global and mission objectives Planning considers avoidance of high speed dynamic cooperative and non-cooperative obstacles
HUMAN-ROBOT INTERACTION		
Robots learn from humans through gesture & speech. Reliable gesture/voice commands for robots that require little training Shared control of tactile procedures using real time sensory feedback Acquire models of un-modelled indoor environments Strongly physically driven (e.g. interaction by touch) Clear differentiation between levels of autonomy Augmented reality-based visual displays	Interact w/users to learn to tackle complex problems Inference of complex intention from natural gesture/voice interactions One controller-multiple robots with acceptable safe standards Robots facilitate simple error recovery HMI uses visual perception (facial & emotion) Basic emotion modelling and interpretation Physical interaction governed by eye-tracking	Companion robots adapt their skills to assist humans Voice and natural language (including translation) Neural (muscle/nerve-ending) & non-invasive brain-based (EEG) interfaces Seamless cooperation interfaces using natural gesture & voice communication Self-arbitrating interfaces for control of multiple robots Natural interfaces that can interpret human intent Modelling of emotion for companion robots/toys Advanced human motion interpretation of unknown/unlearned gestures Human emotion and behaviour interpretation
MISSION PLANNING		
Path and mission planning are performed manually with aid of automated tools Replanning is performed automatically in most domains with minimal interaction Mathematical criteria for predicting the quality of motion and behaviour in planning algorithms	Planning uses extendable knowledge bases prompting user for input Sensor-centric planning and control algorithms capable of operating in reduced information spaces Development of sampling-based planning techniques capable of producing plans in the presence of high dimensionality problems Automatic adaptation of plans and motion when new criteria are introduced Offline realisation of time-optimal planning for situations of high complexity Capacity to sequence planning steps and make adjustments to avoid problems in optimal and sub-optimal fashion under time constraints	Autonomous planning for tasks of high dimensionality & multiple constraints More complex robots (i.e. those with manipulators) plan for themselves Boundaries between sensing, planning, control and learning disappear Theories for unifying and reducing different planning approaches (topological, combinatorial, and dimensional complexities of information spaces) Interleaved planning and execution with model produced online Cognitive approaches to planning available with trade-offs between options Interactive learning – robot can learn or be instructed by human

[17] Cited from Finn (2012, pp. 129-131).

SHORT TERM	MEDIUM TERM	LONG TERM
	SENSING & PERCEPTION	
Data processing through reasoning Simple audio and speech recognition/interaction Capacity to recognise hundreds of objects in real time Object properties retrieved through single sensor modality Capacity to operate in large buildings, shopping malls, etc. Humans recognised and human movements understood Simple human emotion recognised	Capacity to process much higher volumes of data in real time Cheaper sensors with higher resolution, dynamic range and sensitivity Environmental data sensed using complex audio and olfactory sensors Multiple fused sensor modalities (visual/tactile/haptic/auditory/chemical) Known events & activities properly recognised & interpreted in context Object classification of up to 10,000 objects in real time Identification of object properties and affordances Capacity to understand relationships between objects Facial recognition from multiple angles Safety-related classification of scenes Human activity recognised	Integration multiple sensory modalities to acquire models of environment Ability to use them for navigation & interaction with novel objects and events Development of uncertainty management techniques Low cost, high capability sensing systems Perception capabilities approach those of humans Track several hundred objects of interest in real time using multi-modal data Object recognition of in excess of 10,000 objects in real time Reliable extraction of human emotional cues Capacity to interpret human intention Perception affords full autonomy to robots over extended periods
	INTEGRATION & ARCHITECTURES	
Client server architectures, distributed but coupled Layered & hierarchal architectures Vertical integration based on top-down/bottom-up design	Distributed, multi-system architectures Modular multi-vendor sub-system integration Plug-and-play functionality for devices Hybrid architectures with self-describing data and interfaces Better support tools for semi-automatic configuration	Cognitive (i.e. behaviour-based) architectures Distributed architectures (multi-vendor, multi-service, multi-application) Self, configuring (i.e. ‗agnostic‘) architectures Capacity to formally capture & re-use engineering experiences
	GRASPING & MANIPULATION	
Limited pick & place in home & industry Reliable manipulation of non-rigid and non-solid objects Reliably open doors & cabinets (specialised ―hands‖) Different grasping strategies not pre-programmed Grippers and end-effectors become more flexible Grippers with multiple fingers become available	Pre-programmed grasping strategies become obsolete Robust manipulation large, graspable, rigid, & articulated objects/tools without a priori knowledge Grasping strategies accommodate multiple hands Sensors embedded into hands Recognition of object pose changes manipulation strategy Grasping strategies computed in real time based on object orientation and task Grasping skills improve with experience Many different objects can be manipulated, but human dexterity not matched	Nearly human levels of mechanical dexterity Human-like assembly becomes possible Hands covered with high resolution tactile skin Robust, sensor-based prehensile manipulation Limp and highly plastic objects can be handled New technologies for micro/macro grasping and handling Multi-hand cooperation
	ACTUATION & LOCOMOTION	
Standard methods of propulsion (e.g. wheels, tracks, propellers, water jets) Standard methods of locomotion (e.g. engines, fixed/rotary/flapping wings) Biomimetic locomotion (fish mimicking, snakes, indoor bi-pedal walking) Weight reduction, miniaturisation for actuator power/weight ratio Electrical motors with over 20kW direct drive Bi-metallic actuators used in micro applications Ball-and-socket joint assemblies	Special motor concepts (chemical, nuclear, etc.) Smart actuation (SMA, EAP), smart plastics, artificial muscle Energy buffering and saving load potential energy generated by robot Bi-pedal walking on unstructured ground Adhesion for wall-climbing Micro-actuation developed	High performance actuators for lightweight/safe mobile manipulation Capacity to trade response between precision, power and fault tolerance Capacity to draw power from alternate power sources

SHORT TERM	MEDIUM TERM	LONG TERM
COOPERATION & COLLABORATION		
Plan coordination and maintenance conducted by individual robots Distributed planning for multiple robots coordinated with GPS sensors Establishment of ad hoc communication between team and sub-team members Bidding and negotiation systems work effectively in outdoor areas Cooperative perception & navigation with practical planning under uncertainty	Knowledge-based learning using distributed agents Humans start to form a component of the robotic team & vice versa Decentralised planning and decision-making under uncertainty Multiple application (e.g. search and rescue, ISR, etc.) viable	Capacity to cooperate without explicit representation of action Swarm-based reasoning distributed throughout team members Energy autonomy for large teams of robots solved
LEARNING & ADAPTATION		
Robots learn through observation of patterns and trial & error Robots learn from interacting with humans through gesture & speech Acquire models of un-modelled indoor environments Existing reinforced learning techniques are adapted for robotics	Interact w/users to learn to tackle complex problems Robots are able to internalise and use world models Controllers have multiple modules based on learning Learning techniques adapt their behaviour to changing circumstances Robots facilitate simple error recovery	Robots continuously acquire & improve known skills Robots learn from humans and other robots Life-long learning captured and available to robots Learnt behaviour models are available online Companion robots adapt their skills to assist humans Interaction is based on recognition of human intent
SYSTEMS ENGINEERING		
Independent, special-purpose tools for designing robots and sub-systems Simulation of kinematic and dynamic properties to test robotic designs of hardware and software in static environments	More integrated tool-chain exists Systems engineers can easily integrate special purpose tools into the design Simulation of kinematic and dynamic properties of robotics (hardware and software) within its environment becomes easier	Unified framework exists allowing robotic designs to be tailored to application Variety of environmental, kinematic, and dynamic modelling tools exist so that parametric models can easily be modified by designer or end-user Tools allow auto-coding of robotic systems (i.e. software is automatically derived from high-level task descriptions or pre-existing macros) Tools for designing robotic applications focus on the end-user, integration of applications, other systems and web-based services
POWER MANAGEMENT & ENERGY		
Fuel and solar cells, potential energy storage and wired transmission Some elementary harvesting/scavenging from the surroundings (e.g. wave energy, wind energy, bio-ingredients) Mainly chemical, combustion, heat, pneumatic and hydraulic engines Intelligent use of actuators and drives to conserve energy Advanced planning and control algorithms Efficient design of platforms Unused sub-systems disabled Batteries up to 20kW	Wireless power transmission Increased battery energy density	Laser-powered robots Micro-biological power generation

Page classification: UNCLASSIFIED

DEFENCE SCIENCE AND TECHNOLOGY GROUP DOCUMENT CONTROL DATA					1. PRIVACY MARKING / CAVEAT
2. TITLE Trusted Autonomy: Concept Development in Technology Foresight			3. SECURITY CLASSIFICATION Document (U) Title (U) Abstract (U)		
4. AUTHOR Scott Wheeler			5. CORPORATE AUTHOR Defence Science and Technology Group 506 Lorimer St Fishermans Bend Victoria 3207Australia		
6a. DST GROUP NUMBER DST-Group-TR-3153	6b. AR NUMBER AR-016-395		6c. TYPE OF REPORT Technical Report	7. DOCUMENT DATE September 2015	
8. FILE NUMBER 2015/1054142/1	9. TASK NUMBER 07/398	10. TASK SPONSOR US DDR&E	11. NO. OF PAGES 41	12. NO. OF REF. 30	
13. DOWNGRADING/DELIMITING INSTRUCTIONS To be reviewed three years after date of publication			14. RELEASE AUTHORITY Chief, Joint and Operations Analysis Division		
15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT Approved for public release OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SA 5111					
16. DELIBERATE ANNOUNCEMENT No Limitations					
17. CITATION IN OTHER DOCUMENTS Yes					
18. DST GROUP RESEARCH LIBRARY THESAURUS Technology, Autonomous Vehicles, Forecasting					

19. ABSTRACT

This report provides the synopsis of a five year collaborative program of research between DST Group and the University of South Australia in the study of autonomous systems concepts. The purpose of the program is to establish a methodological means of technology foresight, to assess how future technologies shape or contribute to performance in autonomous systems and to identify key technologies of greatest impact. We propose a new model for the categorisation and assessment of autonomy, which provides a systematic and auditable way to explore emerging technologies for autonomous system by means of parametric investigation. Outcomes might then inform decision makers in Defence and National Security to shape the future of policy, strategy, emerging concepts, and force development in autonomous systems and their related technologies.

Page classification: UNCLASSIFIED